Representation and Retrieval Issues in DARWIN:

Digital Analysis and Recognition of Whale Images on a Network

Kelly R. Debure, Eckerd College, St.Petersburg, FL

Background

Identification of individual animals is necessary for the study of behavioral and ecological patterns

- Artificial markings
- Manual photo-identification based on natural markings

Photo-Identification

Dolphin dorsal fins with distinctive natural markings

The DARWIN Project

Goals:

The DARWIN Project

Goals:

- provide access to a database of digital images of dolphin dorsal fins

The DARWIN Project

Goals:

- provide access to a database of digital images of dolphin dorsal fins
- storage and retrieval of observational data which accompanies each fin

The DARWIN Project

Goals:

- provide access to a database of digital images of dolphin dorsal fins
- storage and retrieval of observational data which accompanies each fin
- enable access to images and data across a network

The DARWIN Project

Goals:

- provide access to a database of digital images of dolphin dorsal fins
- storage and retrieval of observational data which accompanies each fin
- enable access to images and data across a network
- reduce the tedium of manual cataloging and access of individual dorsal fins

The DARWIN Project

History:

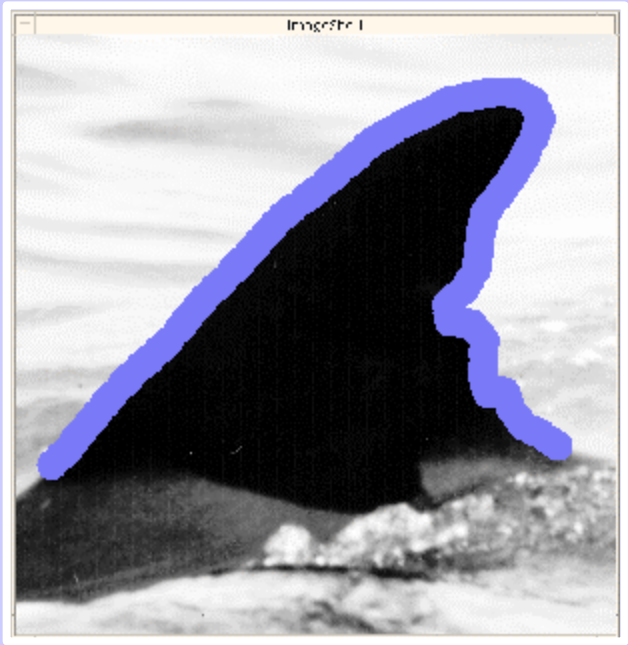
- Initially implemented in 1996 by undergraduate student Mark Allen under the direction of John Stewman
- Originated as and remains a collaborative effort with the Eckerd College Dolphin Research Group headed by John E Reynolds, III
- Originally implemented on a PC under Windows, ported to Unix in 1998 by undergraduate student Dan Wilkin
- Evaluated by Kim Urian, curator of National Marine Fisheries Service Mid-Atlantic Bottlenose Dolphin Photo-ID Catalog

Query Generation

Currently, the user searches for an unknown fin in a database of previously identified fins using a semi-automated sketch based query mechanism.

Query Generation

- user traces a broad outline of the leading and trailing edges of the dorsal fin with the cursor



Query Generation

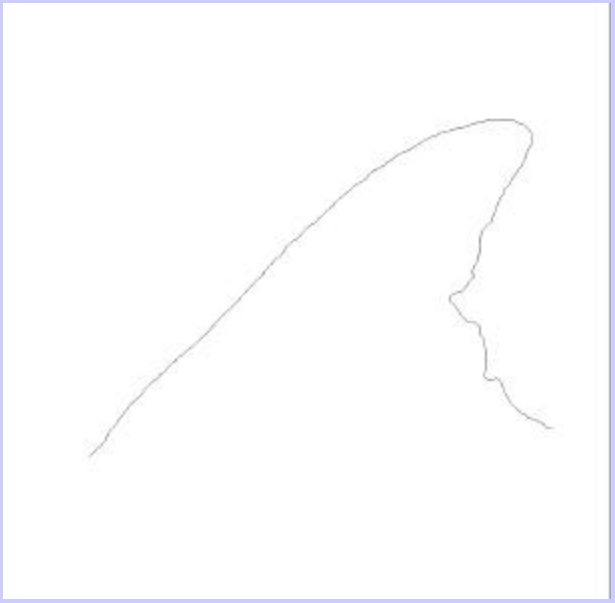
- user traces a broad outline of the leading and trailing edges of the dorsal fin with the cursor
- center of trace initializes the positions of a series of evenly spaced points positioned along the fin edge

Query Generation

- user traces a broad outline of the leading and trailing edges of the dorsal fin with the cursor
- center of trace initializes the positions of a series of evenly spaced points positioned along the fin edge
- active contours move the points from the initial locations to the actual edge of the fin

Query Generation

- user traces a broad outline of the leading and trailing edges of the dorsal fin with the cursor
- center of trace initializes the positions of a series of evenly spaced points positioned along the fin edge
- active contours move the points from the initial locations to the actual edge of the fin
- an outline of the dorsal fin is produced



Query Generation

Currently, the user searches for an unknown fin in a database of previously identified fins using a semi-automated sketch based query mechanism.

The outline is transformed into a representation suitable for efficient comparison with outlines in the database

Query Generation

- query method is somewhat labor intensive
- automate generation of initial outline
- increase sensitivity to deep concave structures

Outline Representation

From a computational standpoint, the ability to reduce the representation of a dorsal fin outline from a sequence of Cartesian coordinates to a representation of distances along fixed distance increments greatly simplifies the comparison of outlines.

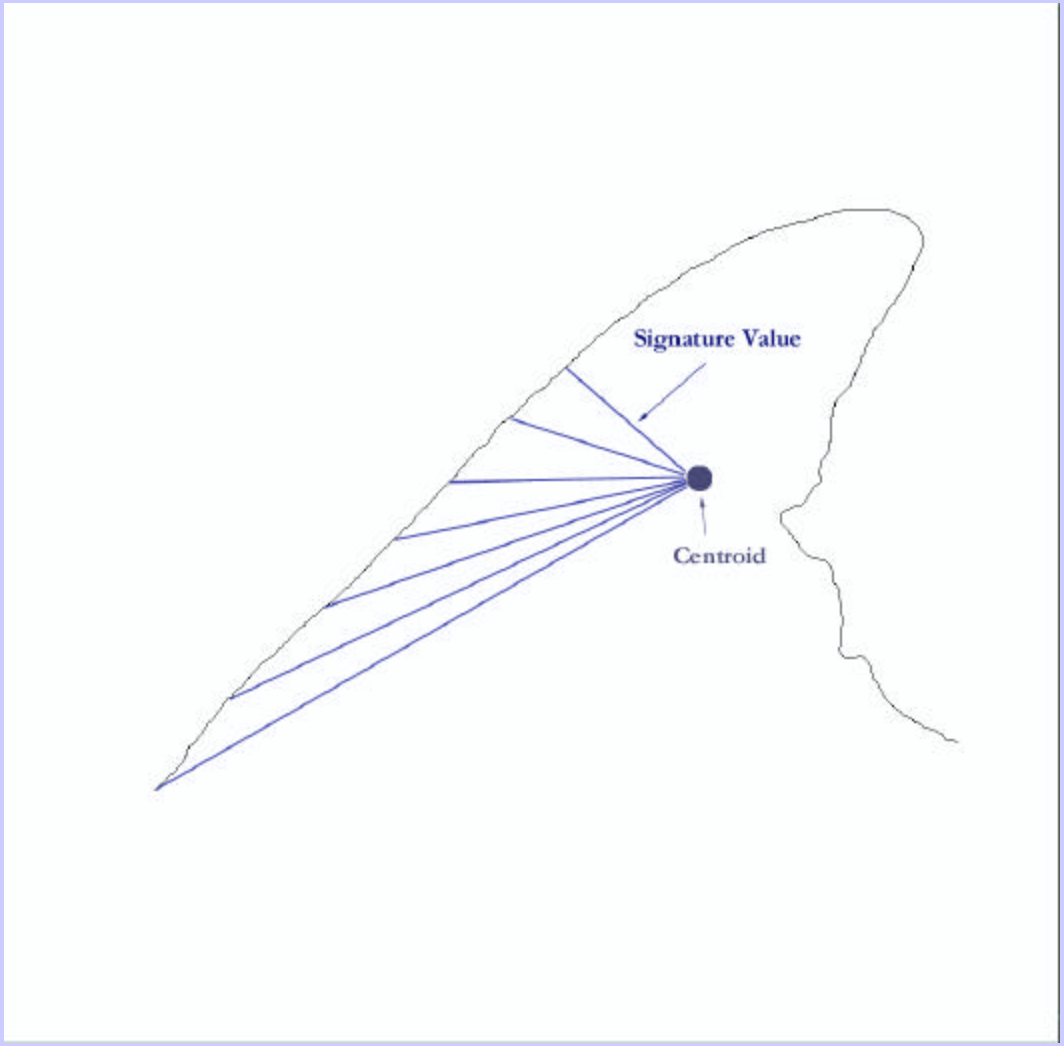
- Reduces the complexity of the computation of the error measure
- Solves some of the problems of comparison, namely translational differences between the fin outlines

Centroid Signature Representation

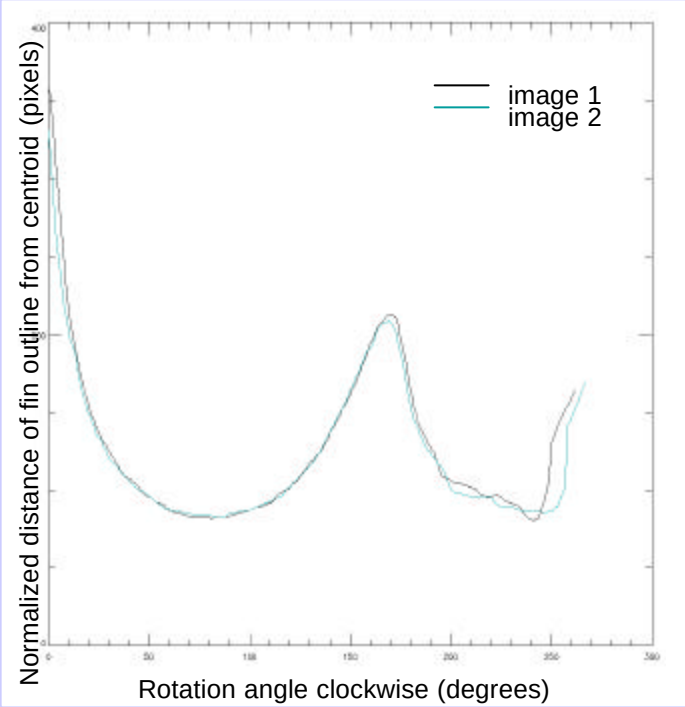
From an outline comprised of Cartesian coordinates a centroid signature is computed as follows:

- Calculate the centroid (c_x,c_y) as the average x and y values of the outline coordinates.
- From the centroid, construct a ray at 0 degrees that intersects the outline of the fin at (x_1,y_1).
- Compute the distance from (c_x,c_y) to (x_1,y_1) and store the distance as the signature value for that angle.
- Increment the angle by one degree, and repeat the process for each of 360 degrees.

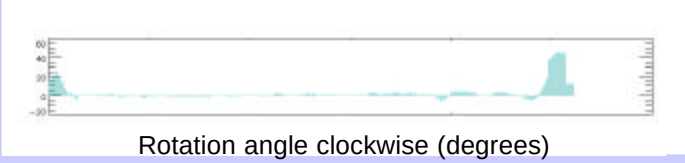
Centroid Signature Representation



Centroid Signature Comparison



Signature error



Original Representation Scheme

Centroid signature representation scheme is effective in reducing the dimensionality of the matching problem. However...

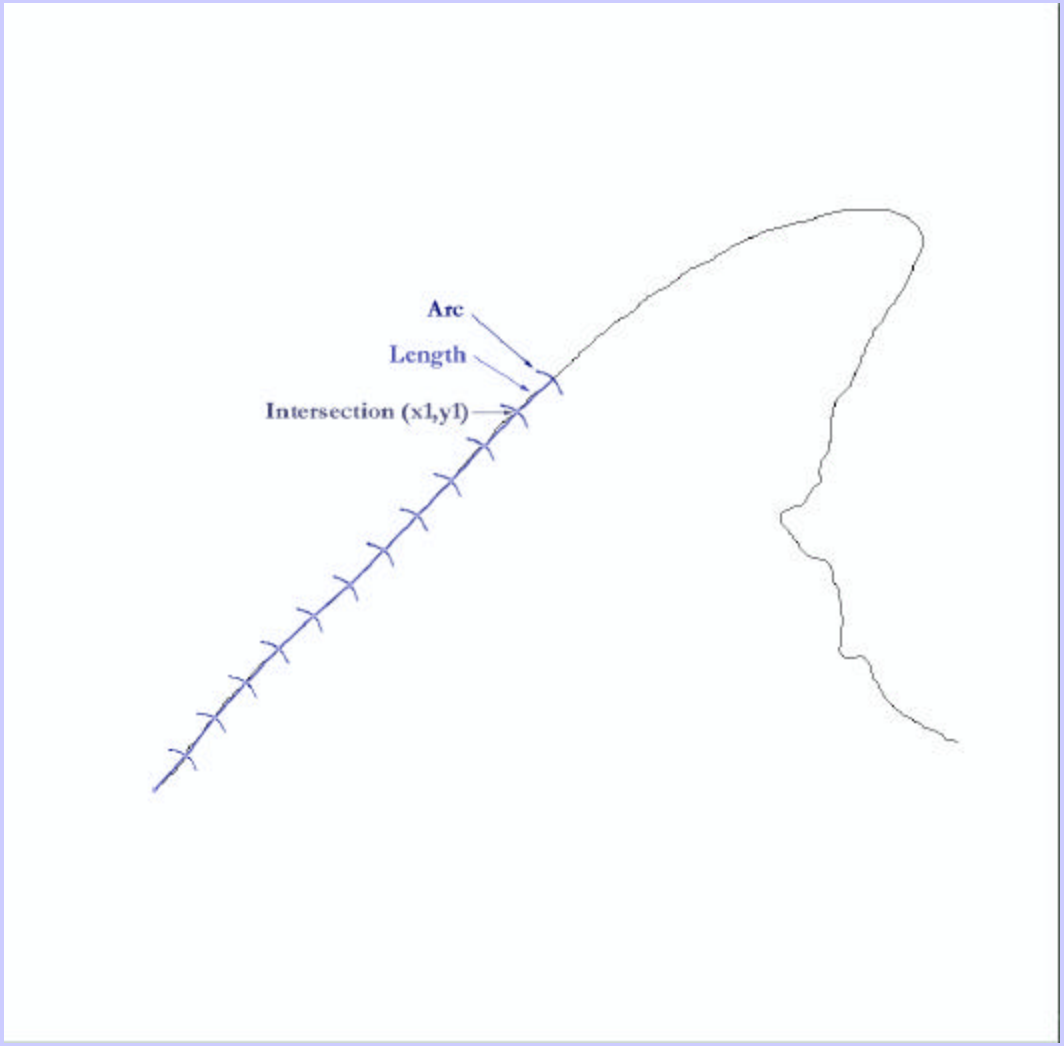
- Match measurement based upon overall fin shape instead of fin damage
- Correction of rotational differences was iterative in nature and computationally intensive.
- Normalization of the average distance from the centroid to the outline to correct scale differences.

Computation of Relative Chain Code

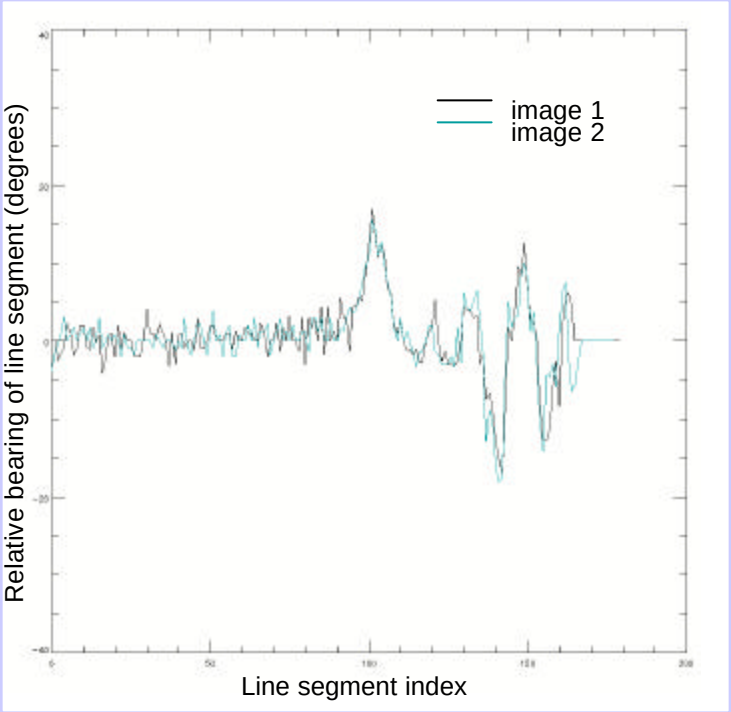
From an outline comprised of Cartesian coordinates the relative chain code is computed as follows:

- Compute the length L of the line segments which result when the outline is partitioned into 180 segments of equal length.
- Compute the point (x_1,y_1) at which an arc of length L, originating at the bottom most point, (x_o, y_o), of the leading edge, intersects the outline.
- Compute the bearing from (x_0, y_0) to (x_1, y_1), and store the difference between the current and previous bearing.
- Repeat the process for each segment

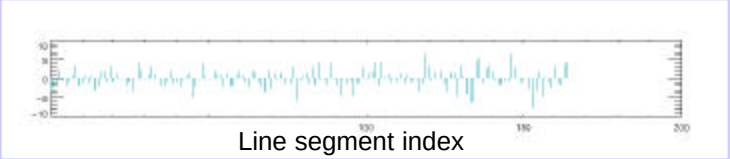
Relative Chain Code Representation



Relative Chain Code Comparison



Signature error



COMPARISON OF MATCH RESULTS *				
method:	execution ** time (min	average rank	ranked top 10%	ranked top 20%
CENTROID BASED SIGNATURE	30.77	19.5	16/30	19/30
RELATIVE CHAIN SIGNATURE	5.05	16.9	19/30	22/30

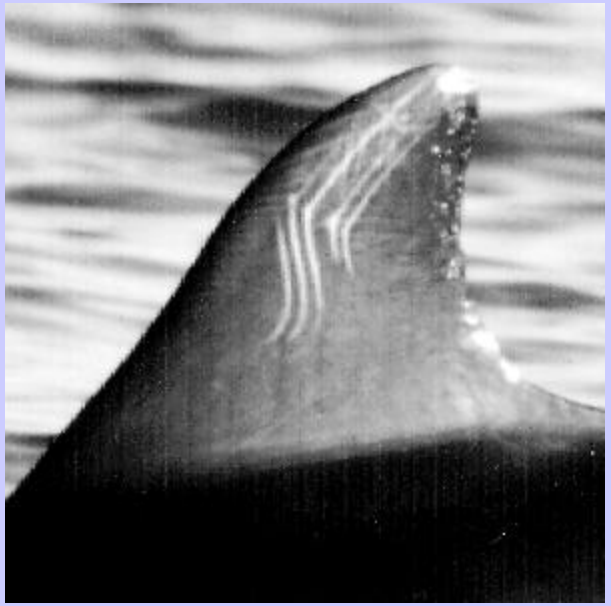
* For preliminary testing, a 100 fin subset of the Eckerd College Dolphin Research Group's 450 fin catalog was scanned and traced to create a test database. Thirty additional images, taken from other photographs or slides of the database fins, were used as "unknown" fins to test the matching performance of the system.

** elapsed time on a Sun Ultra 5

Outline Representation Improvements

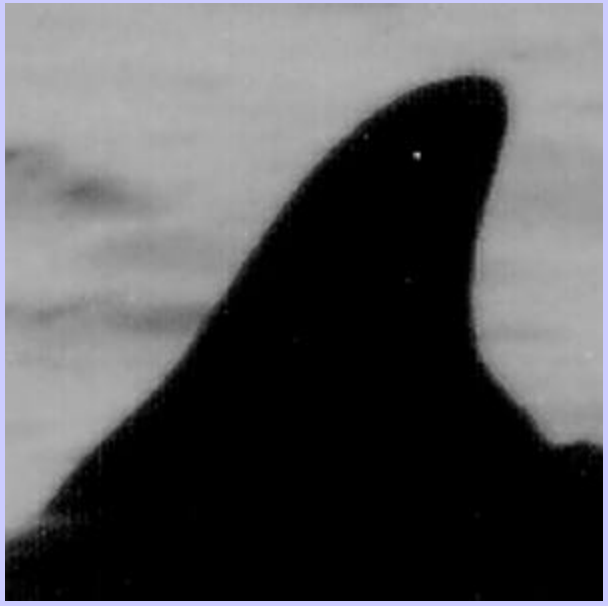
- Distance between key points effectively used in the computation of a scaling factor.
- Rotational differences do not significantly impact the resulting chain values.
- Efficiency/speed greatly improved

Representation Shortfalls



Ignores surface scarring such as rake marks

Representation Shortfalls



Lack of significant notch causes scaling error

Representation Shortfalls

Fails for severely mutilated fins

DARWIN TEAM

Mark Allen
Dan Wilkin
Zach Roberts
John Stewman
Kelly Debure
John E. Reynolds, III

Representation and Retrieval Issues in DARWIN:

Digital Analysis and Recognition of Whale Images on a Network

Kelly R. Debure, Eckerd College, St.Petersburg, FL